Environmental Health Perspectives
Vol. 87, pp. 135–141, 1990

Application of Log-Linear Model in Inference on Karyotypic Evolution in Chronic Myelocytic Leukemia

by Tetsuo Hashimoto,* Megu Ohtaki,* Nanao Kamada,* Hisashi Yamamoto,† and Masaki Munaka*

Relationships among additional chromosome abnormalities in chronic myelocytic leukemia (CML) with translocation 9;22 [Philadelphia chromosome (Ph^1)-positive CML] were analyzed by log-linear models on 709 karyotypes reported in the literature. Additional abnormalities, such as the gain of chromosome 8 (+8), gain of Philadelphia chromosome (+Ph^1), isochromosome of the long arm (q) of chromosome 17 [i(17q)], and the gain of chromosome 19 (+19), were frequently observed. A four-way 2 × 2 × 2 × 2 contingency table was considered with respect to the appearance of these four abnormalities, then the hierarchical log-linear models having at least four main effects were fitted to the observed contingency table. Akaike's information criteria of the models reflected the fitness of the model very well. Parameter estimates of the interaction terms indicated that the combinations of two abnormalities, '+8 and +19', '+Ph^1 and +19', and '+8 and i(17q)' were positively associated, while '+Ph^1 and i(17q)', and '+19 and i(17q)' were negatively associated. Based on the results of the data analysis, an inference was made on the route of karyotypic evolution in Ph^1-positive CML; it statistically supports the hypothesis presented by Heim and Mitelman.

Introduction

The translocation 9;22 found in the karyotypes of chronic myelocytic leukemia (CML) patients is strongly associated with CML as a specific chromosome abnormality. It is regarded as a primary abnormality that plays a fundamental role in initiating the malignant process of CML (*1*). Additional chromosome abnormalities superimposed on the translocation 9;22, such as +Ph^1[+22q-, +del(22)(q11), etc.], +8 and/or i(17q) also occur in most patients with CML in blastic crisis. These abnormalities reflect the karyotypic evolution of malignant cells *in vivo* (*2–4*).

Heim and Mitelman (*4*) presented a hypothesis of karyotypic evolution in CML patients by analyzing the patterns of additional chromosome abnormalities other than translocation 9;22. The major route of karyotypic evolution by their hypothesis was as follows: +8, +Ph^1, or i(17q) are the main additional changes after a translocation 9;22 occurs; +19, on the other hand, seems to occur later in karyotypic evolution, most often in combination with both +8 and +Ph^1; i(17q) in combination with +8 is a quite frequent phenomenon, whereas the combinations 'i(17q) and +19', '+Ph^1 and i(17q)', '+8, i(17q), and +19', and '+Ph^1, i(17q), and +19' are only seen very rarely. These findings testify to the fact that i(17q) apparently has a restrictive role in the cytogenetic evolution in CML, at least when no extra chromosome 8 is present in the cells. The hypothesis, however, is derived from single-column frequencies of combinations of abnormalities, such as '+Ph^1 and +8' and '+8 and i(17q)', without analysis of associations among additional abnormalities.

In the present study, we analyzed 709 cases of Ph^1 positive CML karyotypes with additional abnormalities that were derived from the same database of Mitelman et al. (*4*) and quantified the relationships among additional abnormalities by means of multivariate analysis of frequency table (log-linear models).

Materials and Methods

The material used in this study consisted of 709 Ph^1-positive CML karyotypes with additional abnormalities, which were collected from the *Catalog of Chromosome Aberrations in Cancer*, second edition (*5*). Single-column frequencies of the additional abnormalities were analyzed by our previously reported computer program (*6–8*). A four-way 2 × 2 × 2 × 2 contingency table where the four indices pertain to categorical (none or

*Research Institute for Nuclear Medicine and Biology, Hiroshima University, Kasumi Minami-ku Hiroshima 734, Japan.
†Kure Woman's College, Agaminami Kure 737, Japan.
Address reprint requests to T. Hashimoto, Research Institute for Nuclear Medicine and Biology, Hiroshima University, Kasumi Minami-ku Hiroshima 734, Japan.

present) variables A: +8, B: +Ph^1, C: i(17q), and D: +19 was considered, and then log-linear models (*9*) were fitted to the observed four-way contingency table. The analysis contained 113 hierarchical models with at least four main effects; a saturated model was excluded.

For example, the model only with four main effects, that is, the independent model, is,

$$\log m_{ijkl} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D$$

$[A,B,C,D]$ (Model 1)

where m_{ijk1} is the expected value of observed cell frequency and μ and λ's values are unknown parameters.

The model with four main effects and six terms of two-way interactions, that is, the second-order full model, is:

$$\log m_{ijkl} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{il}^{AD} + \lambda_{jk}^{BC} + \lambda_{jl}^{BD} + \lambda_{kl}^{CD}$$

$[AB,AC,AD,BC,BD,CD]$ (Model 80).

The model with four main effects, six terms of two-way interactions, and four terms of three-way interactions, that is, the third order full model, is:

$$\log m_{ijkl} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{il}^{AD} + \lambda_{jk}^{BC} + \lambda_{jl}^{BD} + \lambda_{kl}^{CD} + \lambda_{ijk}^{ABC} + \lambda_{ijl}^{ABD} + \lambda_{ikl}^{ACD} + \lambda_{jkl}^{BCD}$$

$[ABC,ABD,ACD,BCD]$ (Model 113).

To evaluate the goodness of the fit for 113 models, Akaike's information criteria (AIC) (*10*) was calculated for each model:

$$\text{AIC} = G^2 - 2df$$

where G^2, df represents the likelihood ratio chi-square, the degrees of freedom, of the model, respectively.

Generally, two models M_1 and M_2 are said to be nested if all of the λ effects in M_1 are a subset of the λ values contained in M_2. Conditioning the effects in M_1, the difference in G^2 between the M_1 and the M_2 is the test of the additional effects in M_2. This difference also has an asymptotic chi-square distribution with degrees of freedom equal to the difference in the number of parameters fitted to the two models. Therefore, to test the significance of two- or three-way interactions, the difference in the G^2 between one model and the other model which does not contain the attended interaction was computed. The program package BMDP (*11*) was used for these analyses.

Results and Discussion

Additional chromosome abnormalities frequently found in the 709 karyotypes from Ph^1-positive CML patients were +8 (253 cases, 35.6%), +Ph^1 (238 cases, 33.5%), i(17q) (184 cases, 25.9%), +19 (109 cases, 15.3%), +21 (48 cases, 6.7%), +17 (44 cases, 6.2%), +10 (29 cases, 4.0%), -7 (28 cases, 3.9%), and -Y (17 cases, 2.3%). All of the other additional abnormalities were less than 2% of 709 cases.

Table 1 shows the 2 × 2 × 2 × 2 contingency table with respect to the presence of four main abnormalities (+8, +Ph^1, i(17q), and +19). The table contains the raw data of this analysis.

Table 2 summarizes the degrees of freedom (*df*), likelihood ratio chi-square (LR χ^2), and the AIC of the 113 hierarchical models at least with four main effects. AICs are graphically plotted on the right-hand side of the table. Of all the 113 AICs, 19 AICs were less than 0. The maximum and minimum values of AIC were 188.49 in Model 1 (the independent model) and -5.68 in Model 103. In the models with one or two terms of two-way interactions (Models 2 through 22), all of the AICs were greater than 50. In the models with three terms of two-way interactions (Models 23 through 42), the model $[AD,BC,BD]$ (Model 39) (which has the interactions of '*A*: +8 and *D*: +19', '*B*: +Ph^1 and *C*: i(17q)', and '*B*: +Ph^1 and *D*: +19') indicated the lowest AIC value (11.01). In the models with four terms of two-way interactions (Models 43 through 57), $[AC,AD,BC,BD]$ (Model 53) and $[AD,BC,BD,CD]$ (Model 57) (both of which have the terms, *AD*, *BC*, and *BD* included in Model 39) indicated lower AIC values than other models—that is, 5.65 and 6.36, respectively. These two models have no two-way interaction of '*A*: +8 and *B*: +Ph^1'. Comparison of the AICs of both $[AD,BC,BD]$ (Model 39) and $[AB,AD,BC,BD]$ (Model 49) indicates that the two-way interaction of *AB* does not contribute to increase a fitness. In the models with five terms of two-way interactions (Models 62 through 67), $[AC,AD,BC,BD,CD]$ (Model 67) was the best with an AIC value of −2.93. The AIC was especially lower than other models, and it was the tenth lowest value of all the 113 AICs. The AIC of the second-order full model (Model 80) was −3.69, which was the fifth lowest value. In the models with five terms of two-way interactions and one or two of three-way interactions (Models 81 through 98), Models 89, 92, and 98 (which lacked the two-way term *AB*) indicated especially lower AIC values than others, that is, −1.11, −1.46, and 0.36, re-

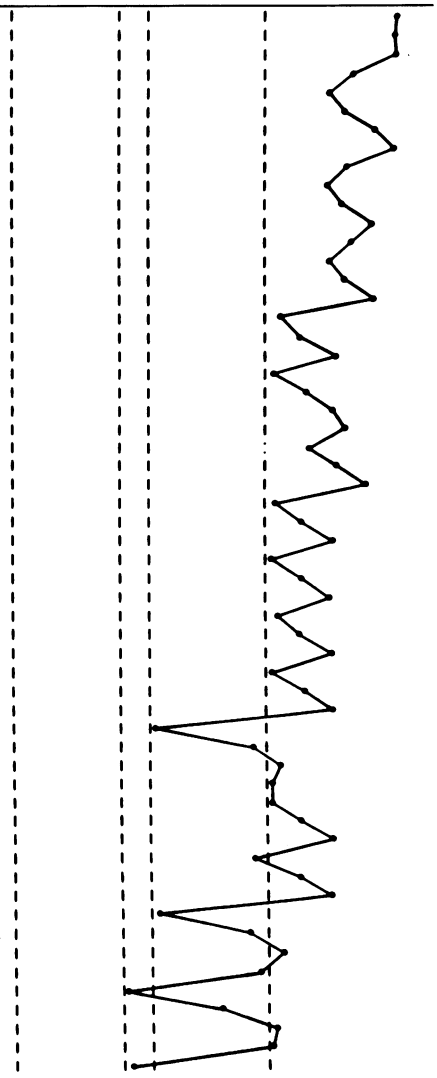
Table 1. Observed frequency table with respect to the presence of four main abnormalities (raw data of the analysis).

Presence of abnormalities			+8 N[a]	+8 P[b]	Total
+19 N	i(17q) N	+Ph^1 N	207	65	272
		+Ph^1 P	106	47	153
		Total	313	112	425
	i(17q) P	+Ph^1 N	102	60	162
		+Ph^1 P	4	9	13
		Total	106	69	175
+19 P	i(17q) N	+Ph^1 N	9	22	31
		+Ph^1 P	27	42	69
		Total	36	64	100
	i(17q) P	+Ph^1 N	1	5	6
		+Ph^1 P	0	3	3
		Total	1	8	9

[a] N: none.
[b] P: present.

Table 2. Degree of freedom (df), likelihood ratio chi-square ($LR\chi^2$), and Akaike's information criteria (AIC) of the 113 hierarchical models at least with four main effects.

Model no.[a]	(A,B,C,D +)		df	$LR\chi^2$	AIC
1		[A,B,C,D]	11	210.49	188.49
2	AB	[C,D,AB]	10	203.40	183.40
3	AC	[B,D,AC]	10	206.32	186.32
4	AD	[B,C,AD]	10	161.37	141.37
5	BC	[A,D,BC]	10	133.20	113.20
6	BD	[A,C,BD]	10	153.42	133.42
7	CD	[A,B,CD]	10	188.25	168.25
8	AB AC	[D,AB,AC]	9	199.23	181.23
9	AB AD	[C,AB,AD]	9	154.27	136.27
10	AB BC	[D,AB,BC]	9	126.11	108.11
11	AB BD	[C,AB,BD]	9	146.33	128.33
12	AB CD	[AB,CD]	9	181.16	163.16
13	AC AD	[B,AC,AD]	9	157.20	139.20
14	AC BC	[D,AC,BC]	9	129.03	111.03
15	AC BD	[AC,BD]	9	149.25	131.25
16	AC CD	[B,AC,CD]	9	184.08	166.08
17	AD BC	[AD,BC]	9	84.08	66.08
18	AD BD	[C,AD,BD]	9	104.29	86.29
19	AD CD	[B,AD,CD]	9	139.12	121.12
20	BC BD	[A,BC,BD]	9	76.13	58.13
21	BC CD	[A,BC,CD]	9	110.96	92.96
22	BD CD	[A,BD,CD]	9	131.18	113.18
23	AB AC AD	[AB,AC,AD]	8	150.10	134.10
24	AB AC BC	[D,AB,AC,BC]	8	116.84	100.84
25	AB AC BD	[AB,AC,BD]	8	142.16	126.16
26	AB AC CD	[AB,AC,CD]	8	176.99	160.99
27	AB AD BC	[AB,AD,BC]	8	76.98	60.98
28	AB AD BD	[AB,AD,BD]	8	103.88	87.88
29	AB AD CD	[AB,AD,CD]	8	132.03	116.03
30	AB BC BD	[AB,BC,BD]	8	69.04	53.04
31	AB BC CD	[AB,BC,CD]	8	103.87	87.87
32	AB BD CD	[AB,BD,CD]	8	124.08	108.08
33	AC AD BC	[AC,AD,BC]	8	79.91	63.91
34	AC AD BD	[AC,AD,BD]	8	100.12	84.12
35	AC AD CD	[B,AC,AD,CD]	8	127.84	111.84
36	AC BC BD	[AC,BC,BD]	8	71.96	55.96
37	AC BC CD	[AC,BC,CD]	8	106.79	90.79
38	AC BD CD	[AC,BD,CD]	8	127.01	111.01
39	AD BC BD	[AD,BC,BD]	8	27.01	11.01
40	AD BC CD	[AD,BC,CD]	8	61.84	45.84
41	AD BD CD	[AD,BD,CD]	8	82.05	66.05
42	BC BD CD	[A,BC,BD,CD]	8	69.48	53.48
43	AB AC AD BC	[AB,AC,AD,BC]	7	67.71	53.71
44	AB AC AD BD	[AB,AC,AD,BD]	7	99.70	85.70
45	AB AC AD CD	[AB,AC,AD,CD]	7	120.74	106.74
46	AB AC BC BD	[AB,AC,BC,BD]	7	59.92	45.92
47	AB AC BC CD	[AB,AC,BC,CD]	7	94.74	80.74
48	AB AC BD CD	[AB,AC,BD,CD]	7	119.36	105.36
49	AB AD BC BD	[AB,AD,BC,BD]	7	26.59	12.59
50	AB AD BC CD	[AB,AD,BC,CD]	7	56.47	42.47
51	AB AD BD CD	[AB,AD,BD,CD]	7	81.64	67.64
52	AB BC BD CD	[AB,BC,BD,CD]	7	62.39	48.39
53	AC AD BC BD	[AC,AD,BC,BD]	7	19.65	5.65
54	AC AD BC CD	[AC,AD,BC,CD]	7	50.55	36.55
55	AC AD BD CD	[AC,AD,BD,CD]	7	70.77	56.77
56	AC BC BD CD	[AC,BC,BD,CD]	7	65.31	51.31
57	AD BC BD CD	[AD,BC,BD,CD]	7	20.36	6.36



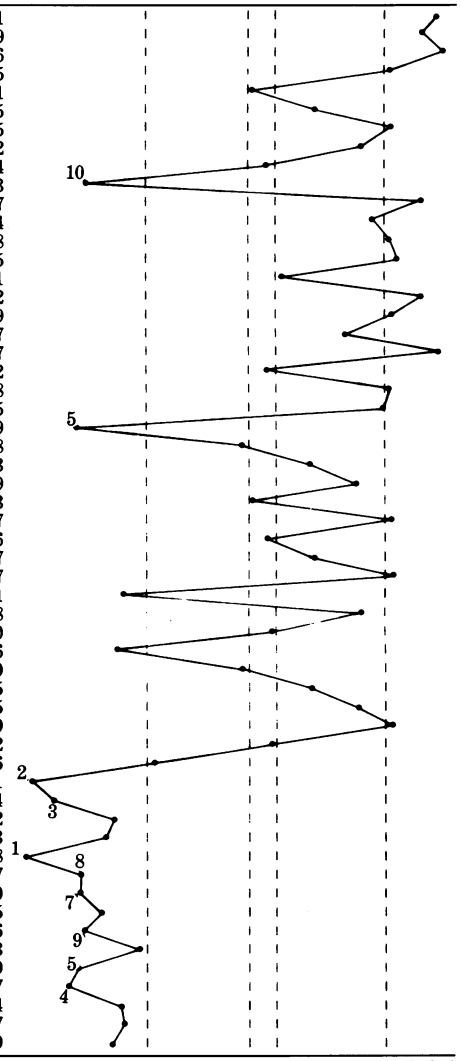
(*Continued on next page*)

spectively. These three AICs, the AIC of [*AC,AD,BC,BD,CD*] (Model 67), and several previous results suggest that the appearance of *A*: +8 and *B*: +Ph[1] tend to be independent. All the models with six terms of two-way interactions and one or more of three-way interactions (Models 99 through 113) had lower AIC values than 0. In these models, the model with six terms of two-way interactions and two of three-way interactions [*CD,ABC,ABD*] (Model 103), indicated a minimum AIC value of all the 113 AICs. The models with six terms of two-way interactions and one of three-way interaction, [*AD,BD,CD,ABC*] (Model 99) and [*AC,BC,CD,ABD*] (Model 100), had the second and the third lowest AICs, respectively.

Table 3 shows the expected frequencies by Models 1, 80, and 103. No good fitness was found between observed frequencies [*O*] and expected frequencies [*E*] by the independent model only with four main effects (Model 1). For example, the value of [*O*]/[*E*] − 1, which indicates the deviation of [*O*] from [*E*], was +3.15 on the cell, corresponding to the presence of '+8, +Ph[1], and +19' (pattern 7). It means that the expected frequency [*E*] was underestimated. Conversely, the value of [*O*]/[*E*] − 1 was −0.74 on the cell of '+19' (pattern 10), meaning that [*E*] was overestimated. Therefore, the presence of +19 may not be explained without considering the interactions of more than two factors. However, the fitness of [*O*] and [*E*] by the second-order full model (Model 80) was considerably improved, and the values of [*O*]/[*E*] − 1 closed with 0. The values of [*O*]/[*E*] − 1 with the best model regarding AIC (Model 103) were nearly 0. The fitness was very good.

Table 2. ***Continued.***

Model no.[a]	(A,B,C,D +)	df	$LR\chi^2$	AIC	Value of AIC[b]
58	AB AC BC ABC	7	114.21	100.21	
59	AB AD BD ABD	7	101.79	87.79	
60	AC AD CD ACD	7	127.66	113.66	
61	BC BD CD BCD	7	68.95	54.95	
62	AB AC AD BC BD	6	17.31	5.31	
63	AB AC AD BC CD	6	38.35	26.35	
64	AB AC AD BD CD	6	70.35	58.35	
65	AB AC BC BD CD	6	53.12	41.12	
66	AB AD BC BD CD	6	19.94	7.94	
67	AC AD BC BD CD	6	9.07	- 2.93	
68	AB AC BC CD ABC	6	91.97	79.97	
69	AB AC BC BD ABC	6	57.14	45.14	
70	AB AC AD BC ABC	6	65.08	53.08	
73	AB AD BD CD ABD	6	79.55	67.55	
72	AB AD BC BD ABD	6	24.51	12.51	
73	AB AC AD BD ABD	6	97.62	85.62	
74	AC AD BD CD ACD	6	70.59	58.59	
75	AC AD BC CD ACD	6	50.37	38.37	
76	AB AC AD CD ACD	6	120.57	108.57	
77	AD BC BD CD BCD	6	19.82	7.82	
78	AC BC BD CD BCD	6	64.78	52.78	
79	AB BC BD CD BCD	6	61.85	49.85	
80	AB AC AD BC BD CD	5	6.31	- 3.69	
81	AB AC AD BC BD ABC	5	14.68	4.68	
82	AB AC AD BC CD ABC	5	35.73	25.73	
83	AB AC BC BD CD ABC	5	50.49	40.49	
84	AB AC AD BC BD ABD	5	15.23	5.23	
85	AB AC AD BD CD ABD	5	68.27	58.27	
86	AB AD BC BD CD ABD	5	17.86	7.86	
87	AB AC AD BC CD ACD	5	38.17	28.17	
88	AB AC AD BD CD ACD	5	70.17	60.17	
89	AC AD BC BD CD ACD	5	8.89	- 1.11	
90	AB AC BC BD CD BCD	5	52.58	42.58	
91	AB AD BC BD CD BCD	5	19.40	9.40	
92	AC AD BC BD CD BCD	5	8.54	- 1.46	
93	AB AC AD BC BD ABC ABD	4	12.60	4.60	
94	AB AC AD BC CD ABC ACD	4	35.55	27.55	
95	AB AC BC BD CD ABC BCD	4	49.95	41.95	
96	AB AC AD BD CD ABD ACD	4	68.09	60.09	
97	AB AD BC BD CD ABD BCD	4	17.32	9.32	
98	AC AD BC BD CD ACD BCD	4	8.36	0.36	
99	AB AC AD BC BD CD ABC	4	2.59	- 5.41	
100	AB AC AD BC BD CD ABD	4	3.46	- 4.54	
101	AB AC AD BC BD CD ACD	4	6.08	- 1.92	
102	AB AC AD BC BD CD BCD	4	5.77	- 2.23	
103	AB AC AD BC BD CD ABC ABD	3	0.32	- 5.68	
104	AB AC AD BC BD CD ABC ACD	3	2.53	- 3.47	
105	AB AC AD BC BD CD ABC BCD	3	2.50	- 3.50	
106	AB AC AD BC BD CD ABD ACD	3	3.45	- 2.55	
107	AB AC AD BC BD CD ABD BCD	3	2.64	- 3.36	
108	AB AC AD BC BD CD ACD BCD	3	5.57	- 0.43	
109	AB AC AD BC BD CD ABC ABD ACD	2	0.31	- 3.69	
110	AB AC AD BC BD CD ABC ABD BCD	2	0.03	- 3.97	
111	AB AC AD BC BD CD ABC ACD BCD	2	2.46	- 1.54	
112	AB AC AD BC BD CD ABD ACD BCD	2	2.63	- 1.37	
113	AB AC AD BC BD CD ABC ABD ACD BCD	1	0.01	- 1.99	



[a] In a hierarchical model, a higher effect cannot be present unless all lower order effects whose indices are subsets of the higher order effect are also included in the model: e.g. if λ^{ABC} is stated (nonzero), it means that μ, λ^{A}, λ^{B}, λ^{C}, λ^{AB}, λ^{AC}, λ^{BC} are all present. All effects other than four main ones are described in the left-hand side of the column. Since only hierarchical models are considered in this analysis, models are also described by a minimal set of effects in parentheses. For example, the full second-order model (Model 80) includes the terms, μ, λ^{A}, λ^{B}, λ^{C}, λ^{D}, λ^{AB}, λ^{AC}, λ^{AD}, λ^{BC}, λ^{BD}, and λ^{CD}, while all the interactions of three or four factors set to 0. This model can be described by the minimal set of effects [*AB,AC,AD,BC,BD,CD*].

[b] Numbers by the symbols (•) represent the order of AIC from the minimum value.

Parameter estimates of the cells that correspond to 'presence', 'presence and presence', or 'presence and presence and presence' of one, two, or three abnormalities were shown in Table 4 as the values of $\hat{\beta} = \exp(\hat{\lambda})$. Concerning the six interactions of a given two factors, if the value of parameter $\hat{\beta}$ is greater than 1, the two factors are considered to be positively correlated to each other. On the other hand, if $\hat{\beta}$ is smaller than 1, a negative correlation is suggested between the two factors. Then the parameter estimates and the results of significance tests on Model 80 indicate that '+8 and i(17q)', '+8 and +19', and '+Ph[1] and +19' were significantly positively associated, while '+22q- and i(17q)' and '+19 and i(17q)' were significantly negatively associated. The parameter estimates of the two-way interactions in the second-order full model (Model 80) was not greatly different from those in the third-order full model (Model 113), and then the parameter estimates of the three-way interactions in the third-order full model (Model 113) and the best model (Model 103) were nearly 1. Therefore, the four-way 2 × 2 × 2 × 2 contingency table is considered to be fully explained by the second-order full model (Model 80). Furthermore, the second-order model, which only lacks a two-way term *AB* (Model 67), is also considered to be a reasonable model.

With regard to the cell of '+8, +Ph[1], +19', the observed frequency was quite higher than the expected value by the independent model (Model 1), as shown in pattern 7 in Table 3. The high frequency of the pattern

Table 3. Observed [*O*] and expected [*E*] frequencies of four-way 2×2×2×2 contingency table.

Pattern number	Abnormalities[a] +8	+Ph[1]	i(17q)	+19	Observed frequency (%) [*O*]	Model 1[b] [*E*]	[*O*]/[*E*]−1	Model 80[c] [*E*]	[*O*]/[*E*]−1	Model 103[d] [*E*]	[*O*]/[*E*]−1
1	−	−	−	−	207(29.1)	189.4	+ 0.09	207.9	− 0.00	207.5	− 0.00
2	−	+	−	−	106(14.9)	96.5	+ 0.09	106.7	− 0.00	106.4	− 0.00
3	−	−	+	−	102(14.3)	67.3	+ 0.51	98.4	+ 0.03	102.5	− 0.00
4	+	−	−	−	65(9.1)	105.8	− 0.38	66.1	− 0.01	66.2	− 0.01
5	+	−	+	−	60(8.4)	37.6	+ 0.59	63.6	− 0.05	59.8	+ 0.00
6	+	+	−	−	47(6.6)	53.9	− 0.12	46.3	+ 0.01	46.9	+ 0.00
7	+	+	−	+	42(5.9)	10.1	+ 3.15	47.2	− 0.11	43.1	− 0.02
8	−	+	−	+	27(3.8)	18.1	+ 0.49	23.8	+ 0.13	27.6	− 0.02
9	+	−	−	+	22(3.1)	19.8	+ 0.11	18.4	+ 0.19	21.8	+ 0.00
10	−	−	−	+	9(1.2)	35.4	− 0.74	12.6	− 0.28	9.5	− 0.05
11	+	+	+	−	9(1.2)	19.2	− 0.53	7.0	+ 0.28	10.1	− 0.10
12	+	−	+	+	5(0.7)	7.0	− 0.28	6.0	− 0.16	6.2	− 0.19
13	−	+	+	−	4(0.5)	34.3	− 0.88	8.0	− 0.50	4.6	− 0.13
14	+	+	+	+	3(0.4)	3.6	− 0.16	2.4	+ 0.25	2.9	+ 0.03
15	−	−	+	+	1(0.1)	12.6	− 0.92	2.0	− 0.50	1.5	− 0.33
16	−	+	+	+	0(0.0)	6.4	− 1.00	0.6	− 1.00	0.4	− 1.00

[a] (+) present; (−) absent.
[b] $\log m_{ijkl} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D$ [*A,B,C,D*].
[c] $\log m_{ijkl} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{il}^{AD} + \lambda_{jk}^{BC} + \lambda_{jl}^{BD} + \lambda_{kl}^{CD}$ [*AB,AC,AD,BC,BD,CD*].
[d] $\log m_{ijkl} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{il}^{AD} + \lambda_{jk}^{BC} + \lambda_{jl}^{BD} + \lambda_{kl}^{CD} + \lambda_{ijk}^{ABC} + \lambda_{ijl}^{ABD}$ [*CD,ABC,ABD*].

Table 4. Parameter estimates [$\hat{\beta} = \exp(\hat{\lambda})$] of log-linear models.

Combination of abnormalities[a]		Model 1[b]	Model 80[c]	Model 113[d]	Model 103[e]
[*A*]	(+8)	0.747	1.065	1.188	1.208
[*B*]	(+Ph[1])	0.714	0.675	0.698	0.676
[*C*]	(i(17q))	0.596	0.395	0.385	0.378
[*D*]	(+19)	0.433	0.380	0.382	0.371
[*AB*]	(+8) × (+Ph[1])		1.081*	1.115	1.124
[*AC*]	(+8) × (i(17q))		1.194*	1.291	1.317
[*AD*]	(+8) × (+19)		1.463*	1.475	1.498
[*BC*]	(+Ph[1]) × (i(17q))		0.630*	0.639	0.617
[*BD*]	(+Ph[1]) × (+19)		1.384*	1.473	1.412
[*CD*]	(i(17q)) × (+19)		0.762*	0.771	0.750
[*ABC*]	(+8) × (+Ph[1]) × (i(17q))			1.125	1.132
[*ABD*]	(+8) × (+Ph[1]) × (+19)			0.909	0.915
[*ACD*]	(+8) × (i(17q)) × (+19)			0.982	
[*BCD*]	(+Ph[1]) × (i(17q)) × (+19)			1.053	
	Value of AIC[f]	188.49	− 3.69	− 1.99	− 5.68
	Order of AIC from the minimum value	113	5	13	1

[a] Factors *A*, *B*, *C*, and *D* represent the categorical (none or present variables, where *A*:+8, *B*:+Ph[1], *C*:i(17q), *D*:+19.
[b] [*A,B,C,D*].
[c] [*AB,AC,AD,BC,BD,CD*].
[d] $\log m_{ijkl} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{il}^{AD} + \lambda_{jk}^{BC} + \lambda_{jl}^{BD} + \lambda_{kl}^{CD} + \lambda_{ijk}^{ABC} + \lambda_{ijl}^{ABD} + \lambda_{ikl}^{ACD} + \lambda_{jkl}^{BCD}$
[*ABC,ABD,ACD,BCD*].
[e] [*CD,ABC,ABD*].
[f] AIC = Akaike's information criteria.
* $p < 0.001$.

is explained only by the interactions of two factors, such as '+8 and +Ph[1]', '+8 and +19', and '+Ph[1] and +19'. The interaction of three factors such as '+8, +Ph[1], and +19' would not be necessary.

Figure 1 schematically shows the route of karyotypic evolution in Ph[1]-positive CML inferred by our analysis. That is, after a translocation 9;22 occurs, one of the four abnormalities, +Ph[1], +8, i(17q), or +19, appears alone at first then the combinations of two abnormalities '+Ph[1] and +8', '+Ph[1] and +19', '+8 and +19', '+8 and i(17q)', '+Ph[1] and i(17q)', or '+19 and i(17q)' appears in the next step. However, the degree of association differ in their combinations: '+Ph[1] and +19', '+8 and +19', and '+8 and i(17q)' were positively associated, while '+Ph[1] and i(17q)' and '+19 and i(17q)' were negatively associated. Furthermore, the third abnormality can be added to the previously formed combination of the two abnormalities (Tables 3,4 and Fig. 1).

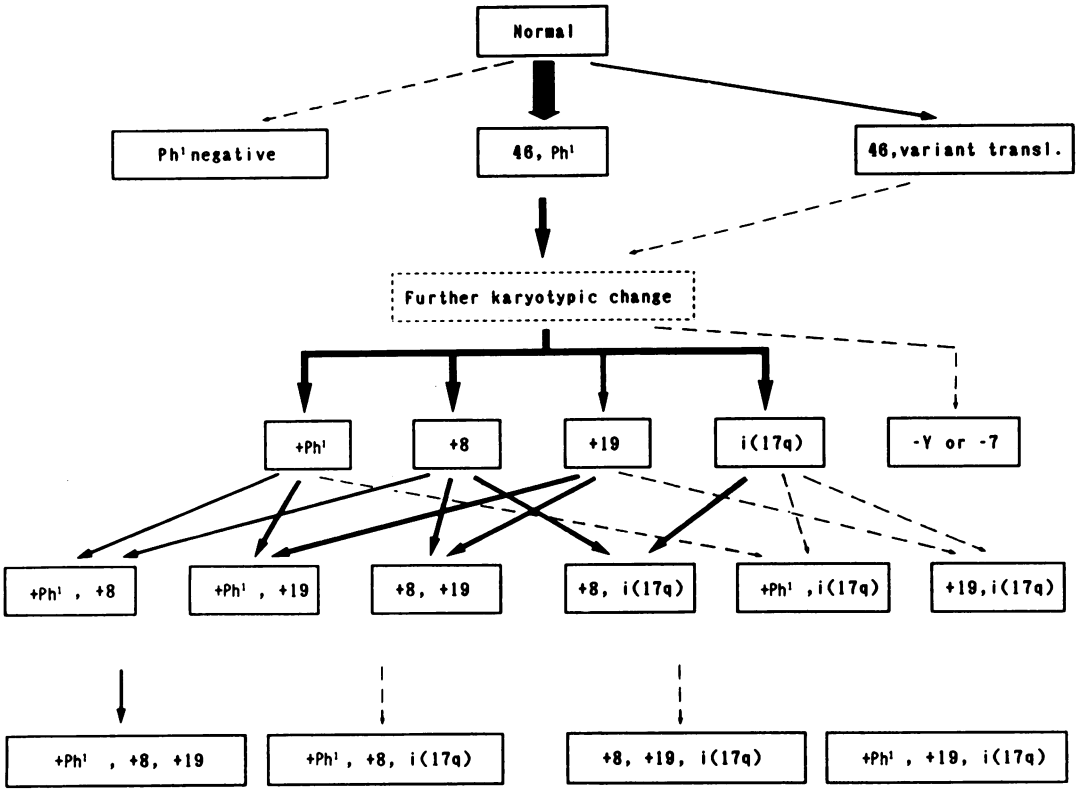


FIGURE 1. Karyotypic evolution in CML. After a translocation 9;22 occurs, one of the four abnormalities, $+Ph^1$, +8, i(17q), or +19 appears alone at first; then the combinations of two abnormalities '$+Ph^1$ and +8', '$+Ph^1$ and +19', '+8 and +19', '+8 and i(17q)', '$+Ph^1$ and i(17q)', or '+19 and i(17q)' appears in the next step. However, the frequencies differ in their combinations, e.g., predominant in '$+Ph^1$ and +19' or '+8 and i(17q)' and rare in '$+Ph^1$ and i(17q)'; furthermore, the third abnormality can be added to the combination of the two abnormalities.

Our findings obtained from analyzing the multiway contingency table with respect to the four additional abnormalities statistically support the hypothesis of karyotypic evolution in CML presented by Heim and Mitelman (*4*). However, our results are derived only from the log-linear analysis. In the next step other methods, such as corresponded analysis or factor analysis, etc., should be applied for the same data set to complement our present analysis. Because the previously mentioned hypothesis results from the analysis based on the karyotypes in a given stage, time-series data analysis of karyotypes from a single patient is also necessary to confirm the hypothesis.

Generally, chromosomal abnormalities of neoplastic cells can be divided into two groups (*1*). One is a primary abnormality that specifically appears in the initiation of the malignant process. In the case of CML, it is a translocation of 9;22. The other is an additional abnormality that appears during the clonal evolution of neoplastic cells, such as +8, $+Ph^1$, and i(17q) in CML. The primary chromosomal changes result in the relocation of cellular oncogenes (*12*). In the initiation of CML, the cellular oncogene c-*abl* mapped on the chromosome 9 band q34, which is normally expressed as 6.0–7.0 kb mRNA, and the protein $P145^{abl}$ is relocated to the *bcr* gene mapped on the chromosome 22 band q11 (*13–15*). This event leads to the expression of rearranged genes and their products, that is 8.5 kb mRNA and protein $P210^{bcr-abl}$, respectively (*16*). The abnormal protein with tyrosine kinase activity would probably make cells acquire a proliferative ability. On the other hand, additional chromosome abnormalities relate to the regulation of proliferation and/or differentiation through loss or inactivation of the genes and unbalance of the genome derived from gene amplification. Thus, additional chromosomal changes contribute to the acceleration of the malignancy of the disease. The analysis of the pattern of additional abnormalities, therefore, would give an important clue to the estimation of pathophysiology and prognosis of CML patients.

REFERENCES

1. Mitelman, F. Restricted number of chromosome regions implicated in aetiology of human cancer and leukemia. Nature 310: 325–327 (1984).
2. Mitelman, F., Levan, G., Nilsson, P., and Brandt, L. Non-random karyotypic evolution in chronic myeloid leukemia. Int. J. Cancer. 18: 24–30 (1976).
3. Heim, S., and Mitelman, F. Secondary chromosome aberrations in the acute leukemias. Cancer Genet. Cytogenet. 22: 331–338 (1986).
4. Heim, S., and Mitelman, F. Multistep cytogenetic scenario in

chronic myeloid leukemia. In: Advances in Viral Oncology, Vol. 7 (G. Klein, Ed.), Raven Press, New York, 1987, pp. 53–76.
5. Catalog of Chromosome Aberrations in Cancer, 2nd ed. (F. Mitelman, Ed.), Alan R. Liss, Inc., New York, 1985.
6. Yamamoto, H., Kamada, N., Tanaka, K., Ueoka, H., Ohtaki, M., Munaka, M., Kuramoto, A., and Ohkita, T. Quantitative analysis of the clinical data on leukemia II. Analysis and rearrangement of human karyotype using computer [in Japanese]. J. Hiroshima Med. Assoc. 34: 1066–1073 (1981).
7. Kamada, N., Yamamoto, H., Tanaka, K., Ohtaki, M., Ueoka, H., Munaka, M., and Kuramoto, A. Analysis and rearrangement of human karyotypes by computer. Cancer Genet. Cytogenet. 10: 17–22 (1982).
8. Hashimoto, T., Kamada, N., Yamamoto, H., and Munaka, M. A computer program for analyses of chromosome abnormalities. Acta Haematol. Jpn. 52: 38–48 (1989).
9. Yanagawa, T. Multivariate Analyses for Discrete Data [in Japanese]. Kyouritsu Shuppan Company Ltd., Tokyou, Japan, 1986, pp. 163–178.
10. Tsuruta, S., and Nogami, Y. AICs in log-linear model for contingency tables with poisson and multinomial designs. J. Jpn. Stat. Soc. 16: 165–172 (1986).
11. Dixon, W. J., and Brown, M. B. BMDP Statistical Software, 1983 revised printing, University of California Press, Los Angeles, CA 1983.
12. Land, H., Parada, L. F., and Weinberg, R. A. Cellular oncogenes and multistep carcinogenesis. Science 222: 771–778 (1983).
13. De Klein, A., Kessel, A. G., Grosveld, G., Bartram, C. R., Hagemeijer, A., Bootsma, D., Spurr, N. K., Heisterkamp, N., Groffen, J., and Stephenson, R. A cellular oncogene is translocated to the Philadelphia chromosome in chronic myelocytic leukemia. Nature 300: 765–767 (1982).
14. Bartram, C., R., De Klein, A., Hagemeijer A., Agthoven, T., Kessel, A. G., Bootsma, D., Grosveld, G., Ferguson-Smith, M. A., Davies, T., Stone, M., Heisterkamp, N., Stephenson, J. R., and Groffen, J. Translocation of c-*abl* oncogene correlates with the presence of a Philadelphia chromosome in chronic myelocytic leukemia. Nature 306: 277–280 (1983).
15. Heisterkamp, N., Stephenson, J. R., Groffen, J., Hansen, P. F., De Klein, A., Bartram, C. R., and Grosveld, G. Localization of the c-*abl* oncogene adjacent to a translocation break point in chronic myelocytic leukemia. Nature 306: 239–242 (1983).
16. Ben-Nerlah, Y., Daley, G. Q., Mes-Masson, A.-M., Witte, O. N., and Baltimore, D. The chronic myelogenous leukemia-specific P210 protein is the product of the *bcr/abl* hybrid gene. Science 231: 212–214 (1986).